

SOCIAL MEDIA MONITORING - GHANA

;

;

© 2021, CIPIT



This work is licensed under the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/legalcode>), which permits unrestricted use, distribution, and reproduction, provided the original work is properly credited.

Cette œuvre est mise à disposition selon les termes de la licence Creative Commons Attribution (<https://creativecommons.org/licenses/by/4.0/legalcode>), qui permet l'utilisation, la distribution et la reproduction sans restriction, pourvu que le mérite de la création originale soit adéquatement reconnu.

IDRC Grant/ Subvention du CRDI: 109109-001-Strengthening cyber policy research centres in the Global South – Centre for Intellectual Property and Information Technology Law

Social Media Monitoring

BEFORE GHANA'S GENERAL ELECTIONS 2020

This report summarizes data, analysis, and findings in the use of social-media (Twitter) influence before Ghana's 2020 General Election.

CIPIT, Strathmore University

Tel +254(0) 703 034 612
Twitter @StrathCIPIT

Madaraka Estate,
Ole Sangale Rd. Nairobi, Kenya.

www.cipit.strathmore.edu
cipit@strathmore.edu

Contents

Executive Summary	1
Data Collection	3
Data Analysis	5
Key Findings and Recommendations	9
About us	17

“Mahama’s campaigns appears to be aggressively using robots and spreading propaganda media”

Executive Summary

As citizens are being more interconnected through social media platforms, the use of disinformation is increasingly becoming a potential weapon against democracy. Disinformation has a fundamental impact on electoral integrity and it has become an issue of global concern¹. Ghana’s pre-election period has experienced many social media interactions, including predictions of the election outcome favouring one candidate or the other, the spread of wrong information (so-called “fake news”), and the sharing of propaganda. This report covers our monitoring, analysis and recommendations.

Analysis and Findings

Our focus has been on Twitter micro-blogging platform, noting that it has a good ratio of young adults to adults, and allows relatively easy identification of trending issues as they happen. Our analysis involved collection of about two hundred and fifty thousand Tweets. The collection of this data considered the following key issues:

- Who are the voters, and who are the political actors in the microblogging platform?
- What are the trending topics, and what or who is driving the conversation?

The conversation about the presidential election surround two main candidates, the incumbent president, Nana Akufo-Addo of the New Patriotic Party (NPP), and the former president John Dramani Mahama of the National Democratic Congress (NDC).

From our data, it is clear that both candidates have employed the use of strategic social media campaigns. There are, however, notable differences. Mahama’s campaigns appears to be aggressively using *robots* (bots) and spreading propaganda media, whereas Akufo-Addo’s campaign relies on authentic (i.e., human-controlled) accounts. This is reminiscent of

¹ “Social Media, Disinformation and Electoral Integrity: IFES Working Paper”, Aug 2019

techniques that *Cambridge Analytica* is believed to have employed to influence elections for their clients².

Encouragingly, we noticed a fact-checking project – Fact-Check Ghana, under the Media Foundation for West Africa – doing an effective job of flagging disinformation shared on the Twitter platform.

Although the efficacy of the strategies of the two campaigns will not be known until well after the election, we can conclude that social media was widely used by both sides.

² African Elections as a Testing Ground: Comparing Coverage of Cambridge Analytica in Nigerian and Kenyan Newspapers <https://doi.org/10.1080/23743670.2019.1679208>

Data Collection

Before the data collection, we identified trending topics on Twitter and collected a sample stream of the Tweets exchanged over the trending topics. We used *Python* and *R* programming languages for tooling to automate this process. Our main approach distinguished the voters and the political actors.

The key guiding questions that help us unveil Twitter campaigns are as follows:

1. Who are the voters, who are the political actors in the microblogging platform?
2. What are the trending topic and who are driving the conversation?
3. Can we establish patterns, motive and the network of the people behind a trending topic?

We collected Twitter data in batches, cleaned and grouped them for analysis. There are three main groups of Ghanaians on Twitter leading the conversations surrounding the upcoming general election:

- A group of Twitter users supporting Nana Akufo-Addo (Pro-Addo group),
- A group that supports John Dramani Mahama (Pro-Mahama), and
- The general group of mixed supporters, media houses and NGOs (General).

After cleaning the collected Tweets and filtering by date ranges (1st Nov – 4th Dec 2020), below is a summary of the data key to this report:

1. For Pro-Addo, 17,126 Tweets were selected based on the hashtags affiliated to Addo's campaigns:

Hashtag	Freq
4More4Nana	384
MaintainNanaAddo	408
TheBattleIsTheLords	463
4MoreToDoMoreForYou	752
VoteNPPToMoveGHForward	1288
NanaIsTheRightChoice	2323
4MoreForNana	6926
VoteNumber1	7822

2. For Pro-Mahama campaigns, we cleaned and used 17342 Tweets with the following hashtags distribution:


Hashtag	Freq
iTrustJM	880
Vote4JohnMahama	1010
JMandJane2020	1061
FaNinyinaa	1195
JohnAndJane2020	1359
KickNanaOut	2212
JohnMahama2020	2452
VoteNumberTwo	3467

3. To capture other general conversations about the election and other political candidates, we collected about 45000 Tweets. After cleaning the repeated Tweets appearing in the above sets, we ended up with about 10000 valid unique tweets. The following hashtags were used in the collection:
- SpecialVoting, LetTheCitizenKnow, RacetotheSeat, DecisionTime,
 - ElectionHQ, GhanaDecides, GhanaElection2020, GhanaElections,
 - Ghana Remembers, Ghana, GhanaFirst, EIB ElectionHub,
 - ElectionCommandCentre, Election360, ElectionHub.

The challenge in data collection is usually with rate limiting from the micro-blogging platform. Even for a paid access, there are limits in place. We constantly monitored the data harvesting tools by reusing some application we had used in the previous Kenya general elections held on 8 August 2017.

App details
Edit

Details and URLs


App icon
Click edit to upload a new icon.

App Name
Elec88

Description
An anti-disinformation and anti-hate speech project. Elec88 collects Tweets from trending hashtags during the Kenyan election period to flag disinformation and hate speech related Tweets.

Website URL

Data Analysis

Summary of analysis done

For each of the three mentioned groups of data sets, we performed the following analysis:

- Timeline Analysis
 - What time are Tweets posted?
- User profiling
 - Who are the top 100 Tweeting account?
 - Are the accounts bot or not?
 - What kind of devices are they using?
- Social network analysis
 - Analyze who are replying to each other,
 - Analyze who are mentioning each other,
 - Find relationship of Retweets among the accounts.
- Geolocation Distribution
 - Where are most of the Tweets coming from?
 - What is the geographic distribution of the accounts?
- Tweet Text
 - Sentiment analysis,
 - Top Favorite Tweets,
 - Top Retweeted Tweets,
 - Top Quoted Tweet,
 - Media count in the Tweets,
 - Most liked media file,
 - Quote proportion,
 - Most used words,
 - Words occurring next to each other,
 - Ratio of verified accounts in the dataset.

Detecting Bots

Twitter Bots can be defined as accounts that can post content or interact with other users in an automated way and without direct human input³. Bots have become a menace in the most widely used social media platforms. They are not easy to detect but there are common characteristics we employ in making our judgment. For instance, the following aspects can be analyzed to determine the likelihood that an account is a bot:

- Profile picture,
- Date of joining a social media platform,
- Ratio of message contribution to message reposting,
- Intervals between making posts or reposting,
- Number of replies made,
- Length of posts, etc.

The biggest challenge in this kind of analysis is the volume of data – thousands of accounts and thousands of tweets to analyze. For our case, we employ the application of machine learning to classify Twitter accounts as bots or not bots in the monitoring process.

Using the aspects listed above, a researcher⁴ has trained a model using both real and fake Twitter accounts to attain 93.53% accuracy when classifying bots and 95.32% accuracy when classifying non-bots. We employed this methodology in the current study.

Detecting Malicious Political Campaigns

Analysis of the collected data allows us to establish patterns. The patterns can, in some cases, reveal a network of different social media operations, and/or can highlight malicious behavior.

After some initial reviews of samples of data, we decided that the inquiries relevant to the Ghana pre-election period are as follows:

- What media is trending?
- Who is/are behind a trending topic?
- Are Twitter bots present and active in trending topics?

³ Bots in the Twittersphere <https://www.pewresearch.org/internet/2018/04/09/bots-in-the-twittersphere/>, Pew Research Center

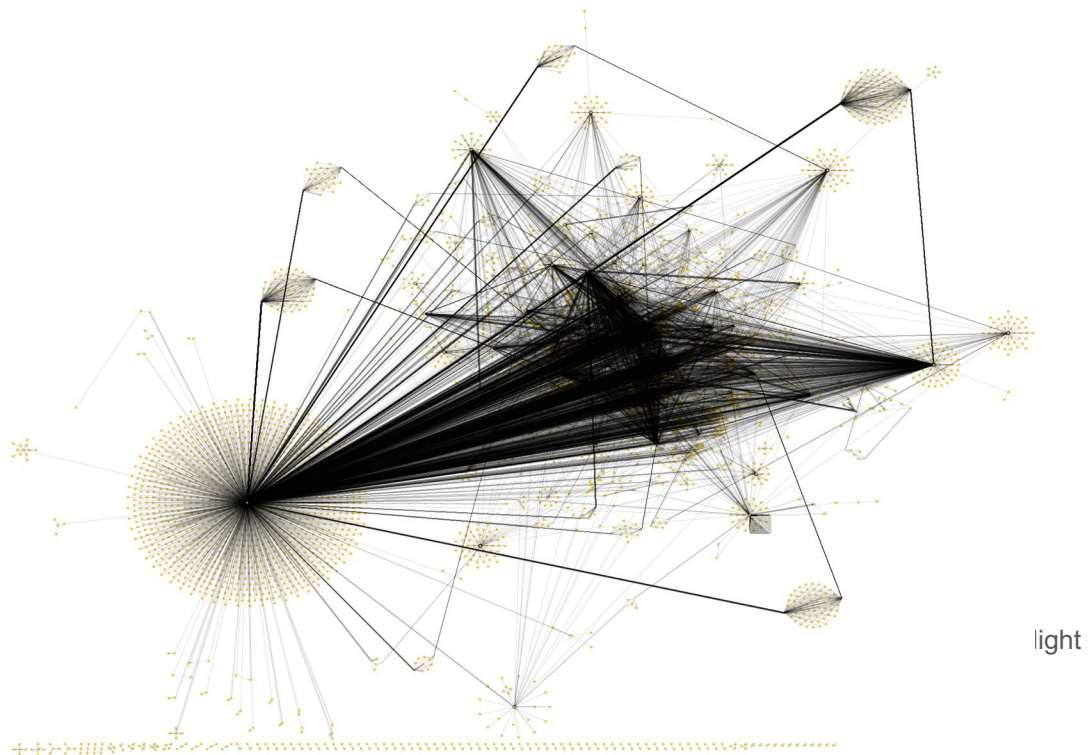
⁴ Michael W. Kearney, Assistant Professor (School of Journalism and School of Informatics Institute) at the University of Missouri

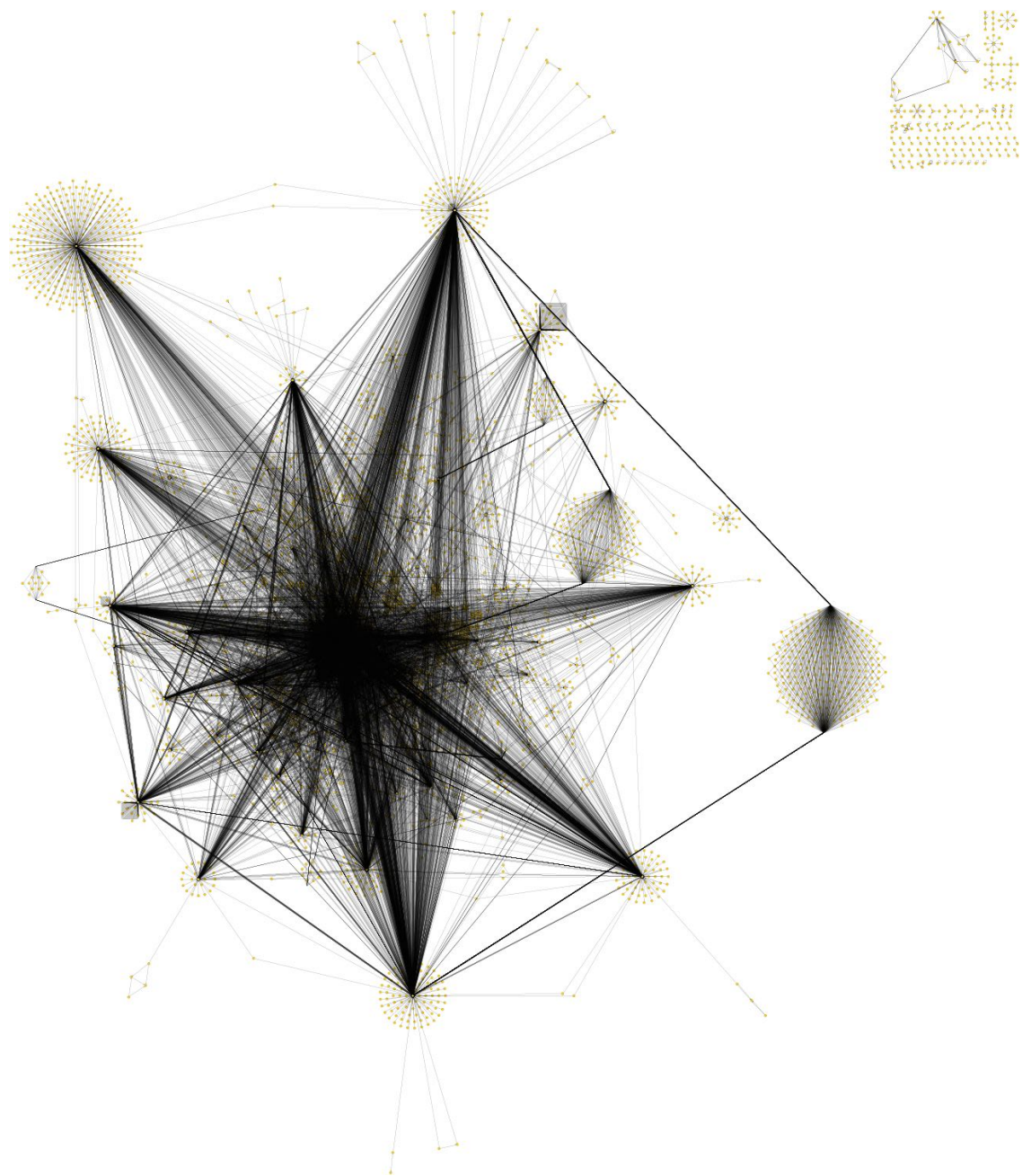
Establishing patterns through visualization

Operating on massive datasets, visualization of the data helps to quickly identify patterns. For the case of Ghana pre-election analysis, we decided to analyse the data in the form of graphs that visually help us pick patterns from the data.

We use an organic layout for the graph generation. A key question for viewing such a graph is whether and where the network has areas of inhomogeneous density. In areas that are denser, more people are connected and there is a higher degree of collaboration and information sharing within the network. This indicator is not necessarily good or bad, but is worth deeper investigation.

Below is an example of an analysis of the accounts mentioned in the pro-Addo's data set. The nodes represent a valid Twitter account. We found out @NAkufoAddo was the key account in most of the mentions.





We found that mentions are scattered in groups. Some social media influencers and marketers were in mention more than Mahama's Twitter account.

Key Findings and Recommendations

Findings

Based on our analysis, Pro-Mahama’s Twitter engagement comprises many Twitter bots (or accounts that behave like bots). Subjecting the top one hundred active accounts in the Bot detection analysis, 25% of the accounts had a positive detection of more than 90 percent accuracy. Three-quarters of the active accounts were treated as malicious.

The top media shared by Pro-Mahama consisted heavily of content flagged as “fake news” by the Fact-Check Ghana project. From our analysis, we established some key accounts behind the spread of the propaganda materials. One of the identified accounts is “Queen Sherifa Gh”



The top media shared by Pro-Mahama consisted of numerous content items flagged as “fake news” by the Fact-Check Ghana project.



The account appears to have been solely created to share propaganda content. The profile image (shown at left) appears to be a random image collected from the internet. We were

able to trace the image to sources with Nigeria⁵ and UK⁶ linkages using Google's reverse image search engine.⁷

Controlling the conversation

We identified the hashtag # MotherSerpentOfCorruption is used to hijack most of the general discussions about the upcoming elections. In fact, we could not accomplish some analysis on the General Dataset because of the overwhelming existence of this hashtag:

Here is one of the persistent error:

```
In addition: Warning messages:
1: Removed 2 rows containing missing values (geom_bar).
2: In wordcloud(word, n, max.words = 100, min.freq = 3, colors = brewer.pal(8, :
  motherserpentofcorruption could not be fit on page. It will not be plotted.
Execution halted
```

Accounts Disappearing

Some Twitter accounts under analysis were either changed or deleted before we could make our conclusions. For instance:

⁵ <https://www.thenicheng.com/how-violence-juju-fuel-rise-in-nigerian-sex-slavery-in-italy/>

⁶ <https://trumpetmediagroup.com/the-trumpet/news/seeking-whistleblowers-and-battling-black-magic-nigeria-ramp/>

⁷ We cannot establish if a Nigerian or UK entity is operating the account, but it is a possibility. Furthermore, we noticed some ads circulating on Twitter for the Pro-Mahama's campaign appeared to share similarities with ads known to have been created by *Cambridge Analytica* (CA) in the Nigerian and Kenyan elections. At this time we cannot conclude that the similarities are sufficient to indicate common origin, but this is an inquiry worth pursuing further.



Our analysis also shows that the accounts use special characters in their profiles to identify each other to either Retweet or Like a post. For instance, from the analysis done, we can confidently state most of the bot-like accounts end with a digit on their screen name, includes Ghana's national flag on their screen name, or have an underscore in their screen name. These are interesting ways to enhance Twitter automation for the Bots or any other valid account that needs to be automated.

Sentiment Analysis

The data sets were subjected to a sentiment analysis to get the overall polarity of the messages conveyed over the Twitter micro-blogging platform. Below summarizes our finding:

General Comments:

Polarity	Count
negative	3278
positive	5024

Pro-Addo:

Polarity	Count
negative	1116
positive	19166

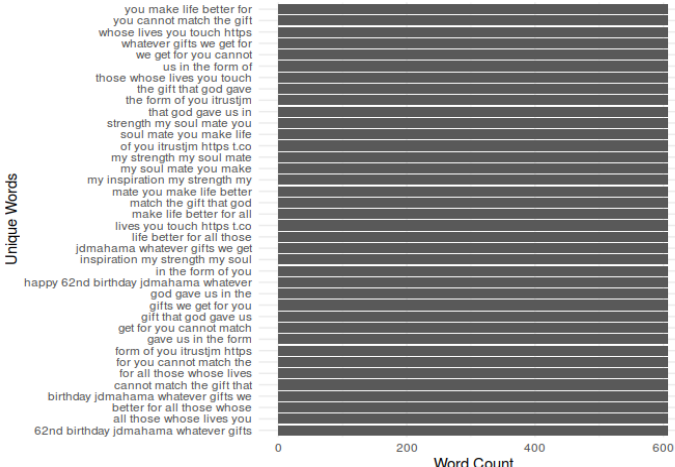
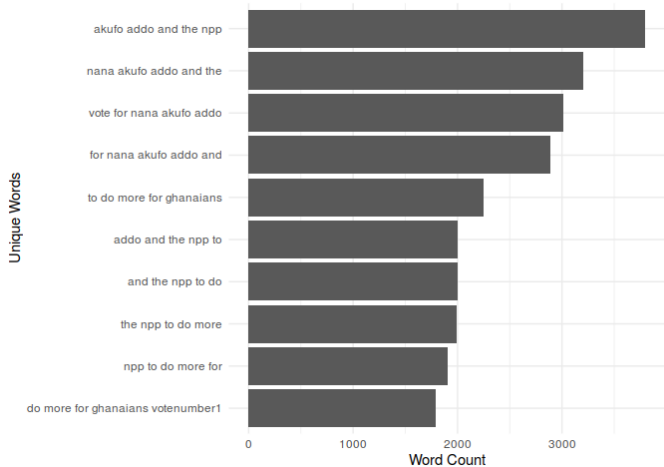
Pro-Mahama:

Polarity	Count
negative	2343
positive	14421

Word Occurrences in Tweets

By analyzing the most repeated words in the data sets, we determined that Pro-Maham’s tweets are repeated in other accounts to a greater extent. This is usually a case where you have a team of Social Media users reading from the same script and mass Tweeting at the same time.

Compared with Pro-Maham’s tweets (right graph, below), Pro-Addo’s Tweets (left graph, below) are more natural with a wider spread of unique words and broader spread of occurrences.

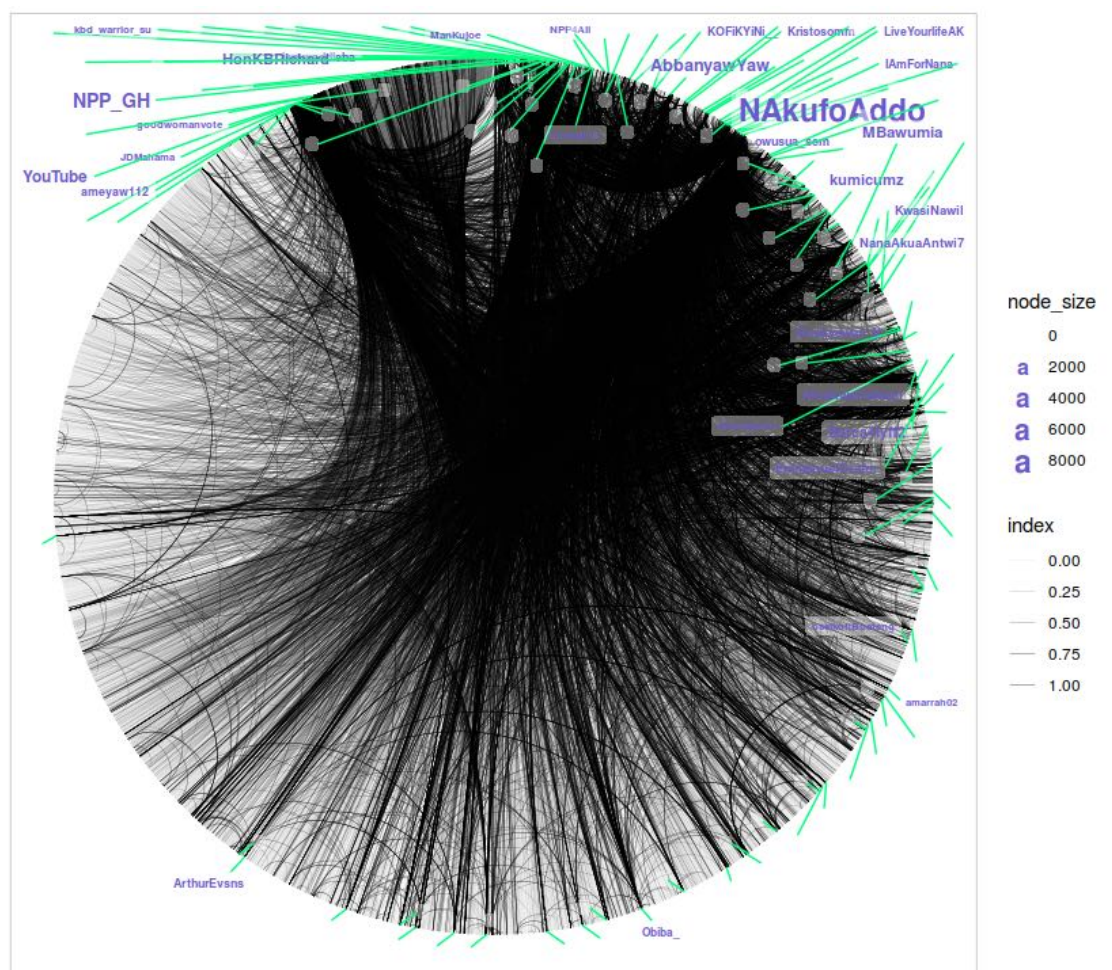


Key Accounts on Twitter

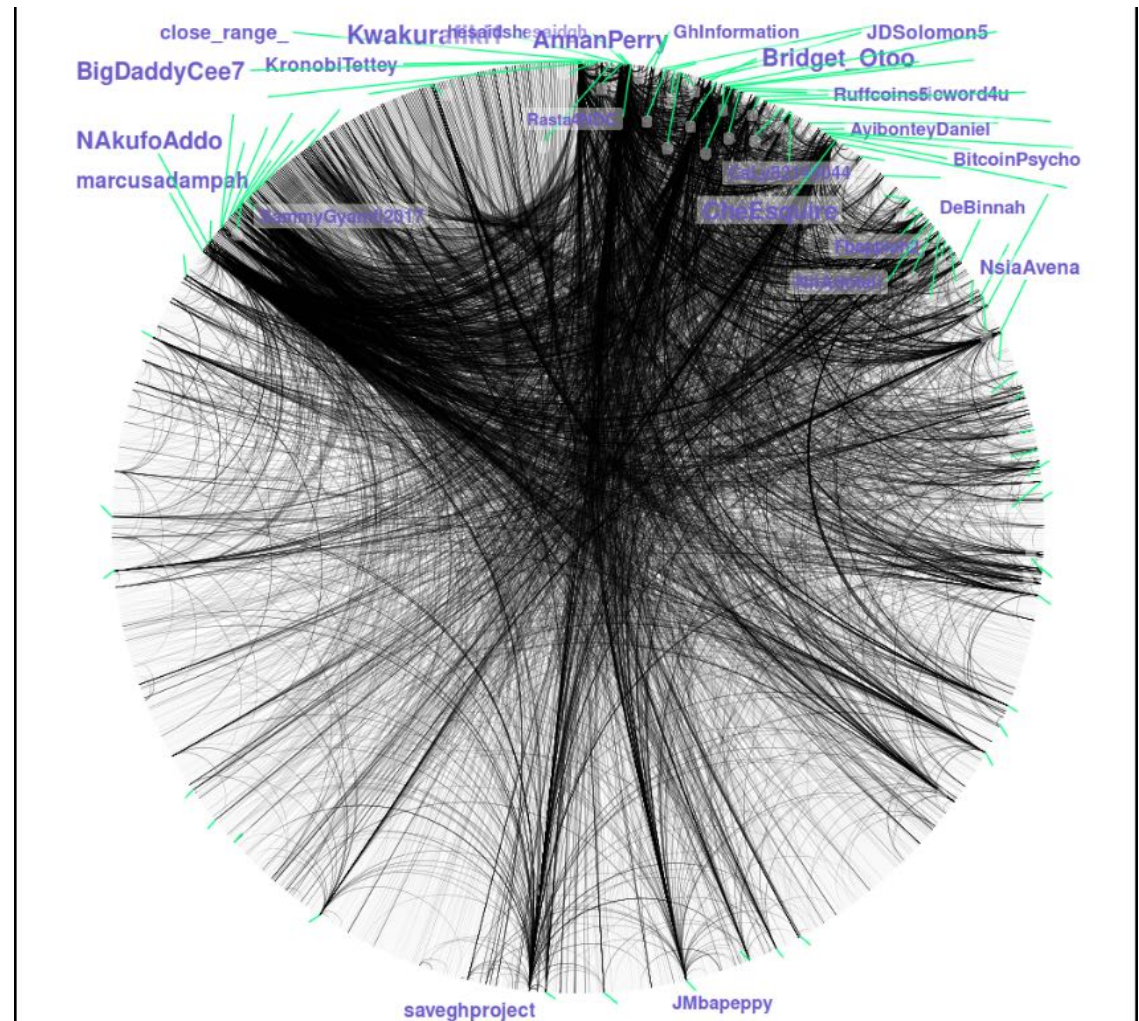
There are certain key accounts dominating conversations or participating in conversations from all the data sets analyzed. Below are the summaries of the accounts visualized with their interactions:

General conversations:

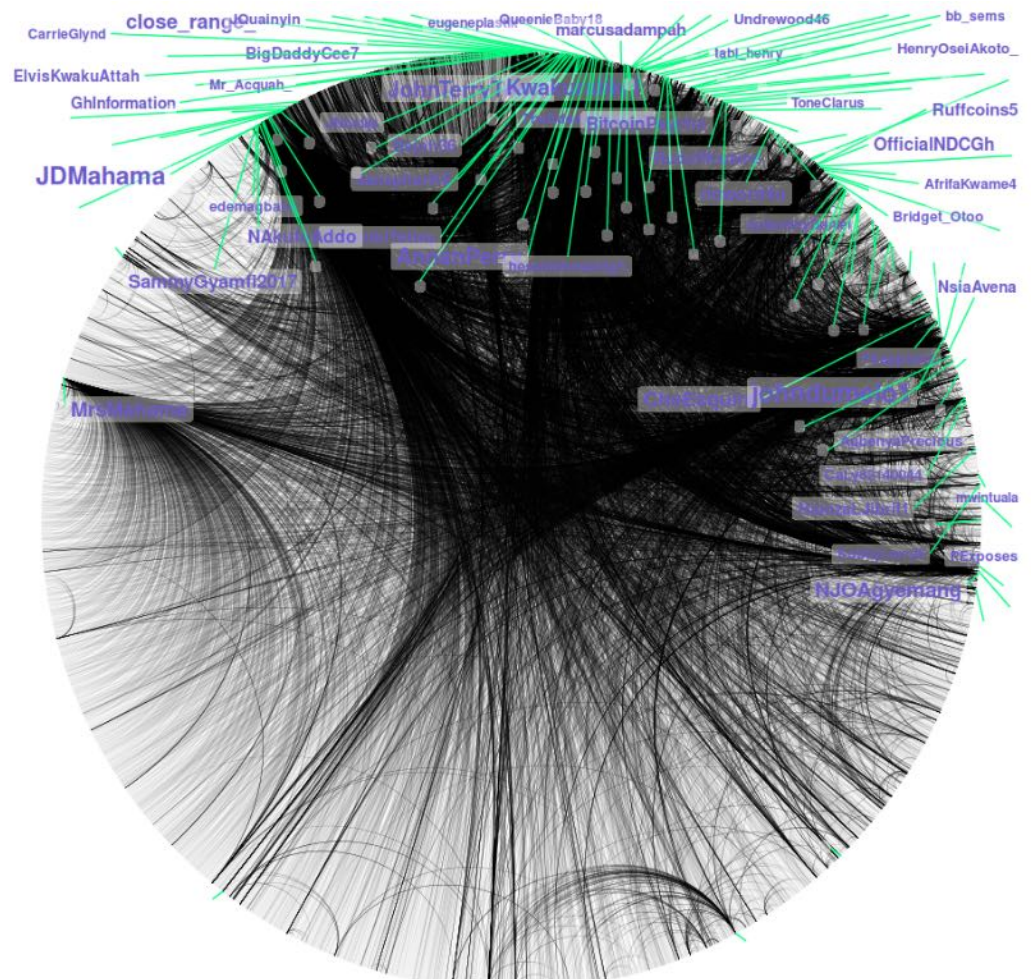
1. Pro-Addo



2. General



3. Pro-Mahama



Recommendations and key vulnerabilities

After analyzing these data sets from our social media monitoring, we encourage the media houses and other researchers to investigate these data further. Further information could help to definitively show that Pro-Mahama campaigners used bots and disinformation to influence the online conversations.

We also recommend further monitoring of social media conversations during and after the election.

Finally, we have noticed that some key Twitter accounts are not verified and have not received attention to gain the authenticity they deserve. For instance, @EUEOMGHANA16 is the official account of election observers from the EU, yet the account is not verified:



Anyone can spoof unverified accounts. There should be an effort to secure this official account so that ill-intentioned individuals are not able to speak on behalf of the EU Election Observation Mission to Ghana. A similar conclusion applies to any other non-verified accounts for official election activities.

About us

CIPIT, Strathmore University

Madaraka Estate,
Ole Sangale Rd. Nairobi, Kenya.

Tel +254(0) 703 034 612

Fax @StrathCIPIT

www.cipit.strathmore.edu